



Funded by
the European Union

 EXPERTISE
FRANCE
GROUPE AFD

Green budget tagging in practice

Lessons and methodological insights from
green budget tagging in Baden-Württemberg,
Lombardy, and Greece

A policy brief for public administrations

May 2026



Funded by
the European Union



Written by: Gintaras Martinaitis, Expertise France capitalisation expert

Funding acknowledgement and disclaimer

The project “Developing Green Budgeting Practices in Baden–Württemberg, Lombardy, and Greece” is funded by the European Union via the Technical Support Instrument and implemented by Expertise France in cooperation with the European Commission.

This policy brief has been produced with the financial assistance of the European Union. The opinions expressed in this document can in no way be taken to reflect the official opinion of the European Union.

Acknowledgements

This policy brief was developed by the Expertise France project team. It draws extensively on the diagnostic analyses, tagging methodologies, and pilot evaluation reports produced, with the support of the project, by the Beneficiary Authorities and their technical partners:

- Ministry of Finance of Baden–Württemberg (Lars Tanzmann and Benjamin Stahl) and FiFo Institute for Public Economics at the University of Cologne (Tom McKenzie and Florian Haerst);
- Ministry of Environment and Climate of Lombardy (Barbara Frascetta) and SDA Bocconi School of Management (Daniela Preite, Silvia Rota and Francesca Casalini);
- Ministry of Economy and Finance of Greece (Effie Tsanti) and ECottrell Ltd. (Jacqueline Cottrell).

The author wishes to thank them for their rigorous technical work, which provided the foundation on which this brief is built, and for their revisions and inputs on the policy brief. Special thanks also go to SG REFORM colleagues for their contribution in the review of the policy brief.

Expertise France is the French interministerial agency in international cooperation and development, part of the French Development Agency (AFD) Group.

Expertise France
Head Office 40, boulevard de Port–Royal – 75005 Paris, France

Links

[expertisefrance.fr](https://www.expertisefrance.fr)
facebook.com/expertisefrance
linkedin.com/company/expertise-france
x.com/expertisefrance
www.youtube.com/@expertisefrance726

© Expertise France, 2026. Reproduction is authorised provided the source is acknowledged.

Contents

Contents	3
List of acronyms and abbreviations	4
Executive summary	5
About this policy brief	6
1. Context: Why green budgeting, and why the tagging framework matters	8
2. Where the Beneficiary Authorities started	9
3. Methodologies developed and tested	10
Step 1: Anchoring in policy priorities	11
Step 2: Defining the scope	13
Step 3: Defining what is "green" (and "brown")	15
Step 4: Assessment procedures	18
Step 5: Establishing governance arrangements	25
4. Lessons and recommendations	27
Annexes	34
Annex A. About the project	34
Annex B. Further reading	36

List of acronyms and abbreviations

AI	Artificial Intelligence
BA	Beneficiary Authority
BW	Baden-Württemberg
CO ₂	Carbon Dioxide
DG ECFIN	Directorate-General for Economic and Financial Affairs of the European Commission
DNSH	Do No Significant Harm
EC	European Commission
EF	Expertise France
EIA	Environmental Impact Assessment
ERDF	European Regional Development Fund
EU	European Union
GBRF	Green Budgeting Reference Framework
GBT	Green Budget Tagging
GHG	Greenhouse Gas
IMC	Interministerial Committee
KlimaG BW	Climate Protection and Climate Change Adaptation Act of Baden-Württemberg
OECD	Organisation for Economic Co-operation and Development
PREAC	Regional Energy, Environment and Climate Programme 2030
PRIA	Regional Plan for Air Quality
PRSS	Regional Programme for Sustainable Development
RRF	Recovery and Resilience Facility
SDG(s)	Sustainable Development Goal(s)
SG REFORM	Reform and Investment Task Force of the European Commission
SIRACC	Regional Integrated Strategy for Climate Change Adaptation
TSI	Technical Support Instrument

Executive summary

Green budgeting has moved rapidly up the EU policy agenda, and many administrations are now moving from commitment to implementation. As this happens, the need is becoming more practical: administrations require tested tagging frameworks, usable tools and institutional arrangements that connect green budgeting with day-to-day budget processes.

This policy brief draws on the EU-funded project “Developing Green Budgeting Practices in Baden-Württemberg, Lombardy, and Greece”. The project supported three administrations at different stages of development: Lombardy built a tagging methodology from scratch; Baden-Württemberg developed and piloted a methodology that extends an existing climate and environmental tagging system, which already identified positive contributions, to cover climate-neutral and climate-harmful spending; and Greece received diagnostic analysis and recommendations for strengthening an established national system. Together, these cases show that there is no single model for green budget tagging (GBT). The tagging framework has to fit the administration’s legal system, budget process, data availability, institutional culture and intended use of results.

The project’s experience points to five core design steps: anchoring tagging in policy priorities; defining the scope; defining what counts as “green” and “brown”; designing assessment procedures; and establishing governance. Across these steps, the most important finding is that methodological choices are real choices, not technical defaults. Starting with a focused climate scope can make implementation more manageable, while full coverage of the EU Taxonomy can avoid future rework. Tagging at commitment – decree level or at budget – item level – can both be credible, provided the level of granularity is feasible and the people with the necessary contextual knowledge are involved.

The project also showed that the main implementation challenges are not purely technical. Negative tagging is politically sensitive and requires careful framing. Line managers need to be involved because they understand the spending, but central quality assurance is needed to ensure consistency. Piloting, training and practical tools are essential because methodologies that look complete on paper often need adjustment when applied to real budget items. Tagging results also need a route into budget decisions – through budget instructions, new funding requests, dashboards, parliamentary documents or management discussions – otherwise green budgeting risks remaining a reporting exercise.

The brief therefore recommends that administrations anchor green budgeting in recognised commitments, calibrate scope deliberately, treat early implementation as a learning phase, invest in tools and training, choose a sufficiently granular tagging unit, handle negative and mixed effects credibly, combine decentralised tagging with central oversight, and connect tagging to budget decisions from the outset.

The central message is simple: green budget tagging is not a one-off technical exercise. It is a multi-cycle reform process that depends on methodology, governance, information quality, communication and institutional commitment moving together.

About this policy brief

Purpose

Green budgeting has moved rapidly up the EU policy agenda. But a persistent gap remains between policy ambition and practice. Top challenges include a lack of technical expertise, fitting methodologies, and staff capacity.¹

This policy brief responds to that need. It draws on the project “Developing Green Budgeting Practices in Baden-Württemberg, Lombardy, and Greece” – the green budget tagging methodologies developed for these regions, and the technical recommendations for Greece’s existing methodology. Rather than offering abstract principles, it presents concrete approaches tested in real budgetary contexts.

Green budgeting includes several tools, such as ex-ante environmental impact assessments and ex-post evaluations. This policy brief focuses on green budget tagging – the most widely used of these tools across EU Member States.

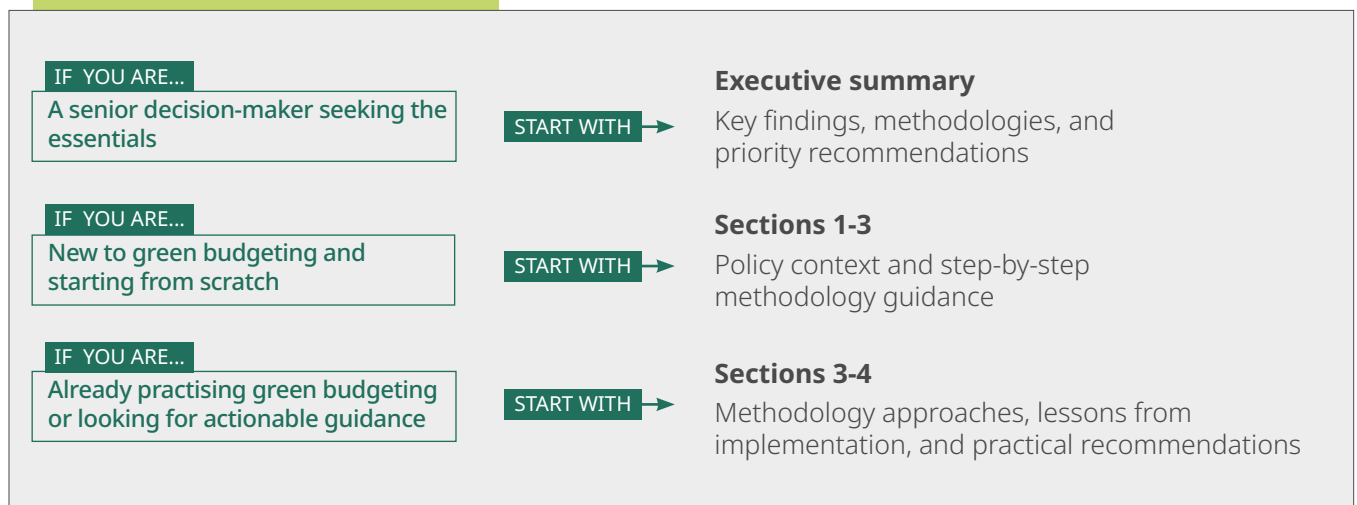
Audience

This brief is intended primarily for:

- Policymakers in national and subnational administrations considering whether and how to introduce or improve green budgeting.
- Budget officials responsible for designing or operating tagging methodologies.
- Environmental policy staff working to integrate climate and environmental objectives into public financial management.

It may also be useful for EU institutions, international organisations, and researchers interested in how green budget tagging works in practice.

How to read this brief



¹ European Commission (2025). Green Budgeting in the EU: Key Insights from the 2025 European Commission Survey on Green Budgeting

Section 3, which walks through the five steps of tagging framework development, uses a single budget item – an incentive scheme for low-emission biomass heating – as a running example. Following this example through the section shows how abstract methodology choices translate into concrete tagging decisions.

What this brief is - and is not

This brief complements, but does not replace, the EU Green Budgeting Reference Framework (GBRF).² The GBRF provides the overarching toolkit for Member States; this brief offers practical experience from administrations that have applied it. The tagging frameworks presented here were developed for specific contexts: a large German federal state (Baden-Württemberg), a major Italian Region (Lombardy), and the Greek national government. They reflect choices shaped by their respective governance contexts, existing systems, and political constraints. Other administrations will need to make their own choices. But the questions each methodology answers – and the trade-offs involved – are broadly transferable.

² <https://economy-finance.ec.europa.eu/system/files/2023-05/European%20Union%20Green%20Budgeting%20Reference%20Framework.pdf>

1. Context: why green budgeting, and why the tagging framework matters

The European Green Deal commits the EU to net-zero emissions by 2050.⁴ Reaching the 2030 target of a 55% reduction requires an estimated €477 billion in additional annual investment.⁵ With Recovery and Resilience Facility funding ending, a gap of roughly €180 billion will open for 2024–2030.⁶ Bridging it depends in part on aligning public budgets – at all levels of government – with climate and environmental goals.

Green budgeting is one of the tools for doing so. It identifies whether budgetary items contribute positively or negatively to environmental goals, helping governments identify where public money supports green priorities and where it underwrites harmful activity – information they can use to inform future budget decisions. The need extends beyond national governments: subnational authorities control on average 40% of all public spending and 55% of all public investment across the OECD.⁷ They also often have direct jurisdiction over housing, transport, land use, and waste – areas central to the green transition.

EU policy has moved quickly to support this agenda. The EU Green Budgeting Reference Framework (GBRF)⁹, developed jointly by the Commission and Member States, provides the overarching toolkit. The EU Taxonomy¹⁰ offers a common classification for environmentally sustainable activities. Technical support has come through TSI projects, biennial surveys, and annual conferences.¹¹ According to the 2025 European Commission Survey, 21 of 26 reporting Member States now have or plan to introduce some form of green budgeting – up from 17 in 2023.¹² Fifteen Member States now practise green budget tagging, with most covering all six EU Taxonomy objectives.

Yet adoption is running ahead of implementation quality. The 2025 Survey identifies the most common

Green budgeting

A process whereby the environmental contributions of budgetary items are identified and assessed with respect to specific performance indicators, with the objective of better aligning budgetary policies with environmental goals (i.e., the contribution of budgetary items to green objectives). (European Commission)³

Green budget tagging

Green budget tagging involves systematically assessing individual budget measures and classifying them according to whether they support or undermine climate and environmental objectives – including climate change, biodiversity, air and water quality, and pollution. Measures are typically tagged as positive (“green”), negative (“brown”), neutral, or mixed. (Adapted from OECD).⁸ The terms “green” and “brown” are used to distinguish environmentally and climate-friendly activities or expenditures (“green”) from environmentally harmful or emissions-intensive ones (“brown”).

³ Pojar, S. (2023). How green budgeting is embedded in national budget processes. European Commission, Directorate-General for Economic and Financial Affairs. Publications Office of the European Union. <https://doi.org/10.2765/285408>

⁴ European Commission (2019). The European Green Deal. COM(2019) 640 final.

⁵ https://single-market-economy.ec.europa.eu/system/files/2023-03/SWD_2023_68_F1_STAFF_WORKING_PAPER_EN_V4_P1_2629849.PDF

⁶ Pisani-Ferry, J., Tagliapietra, S. and Zachmann, G. (2023), cited in Pisani-Ferry, J., Tagliapietra, S. (2026), “An investment strategy to keep the European Green Deal on track,” Bruegel Policy Brief, January 2026. <https://www.bruegel.org/policy-brief/investment-strategy-keep-european-green-deal-track>

⁷ <https://www.oecd.org/en/topics/subnational-finance-and-investment.html>

⁸ OECD (2021), Green Budget Tagging: Introductory Guidance & Principles, OECD Publishing, Paris, <https://doi.org/10.1787/fe7bfcc4-en>

⁹ <https://economy-finance.ec.europa.eu/system/files/2023-05/European%20Union%20Green%20Budgeting%20Reference%20Framework.pdf>

¹⁰ https://finance.ec.europa.eu/sustainable-finance/tools-and-standards/eu-taxonomy-sustainable-activities_en

¹¹ https://economy-finance.ec.europa.eu/economic-governance-framework/national-fiscal-frameworks-eu-member-states/green-budgeting-eu_en

¹² European Commission (2023). Green Budgeting in the EU: Key Insights from the 2023 European Commission Survey on Green Budgeting Practices

¹³ European Commission (2025). Green Budgeting in the EU: Key Insights from the 2025 European Commission Survey on Green Budgeting

barriers as a lack of technical expertise, fitting methodologies, and staff capacity. Sixty percent of practitioners report no clear effects of green budgeting on policy development or resource allocation.¹³ The challenge has shifted from whether to adopt green budgeting to how to make it work. And the “how” turns primarily on the design of the tagging framework, encompassing the methodology, governance, deliverables and connection to the budget process, which together determine whether classification is consistent, whether results are credible, and whether the information reaches decision-makers at the right time and of sufficient quality to inform budgetary decisions.

Methodology is not the only thing that must be clear; so must the purpose of tagging. Green budget tagging is one tool among several, alongside regulations, expenditure reviews, and environmental impact assessments. Its distinctive value lies in becoming a routine part of the budget process – tracking how public money aligns with green objectives, raising visibility, and improving accountability. These benefits only materialise when tagging informs decisions rather than serving as a reporting exercise.

These challenges are particularly acute at the subnational level, where resources and expertise might be more limited. The project was designed to address this gap. Funded by the EU via the Technical Support Instrument and implemented by Expertise France, the project supported three administrations: Baden-Württemberg (Germany) and Lombardy (Italy) at the subnational level, and Greece at the national level. Each developed or strengthened a tagging methodology tailored to its context. The project combined methodology development with capacity building and peer exchange, producing approaches that other administrations can adapt.

2. Where the Beneficiary Authorities started¹⁴

The project supported three Beneficiary Authorities from different starting points in green budget tagging. Understanding their starting points – and the challenges identified through diagnostic analysis – provides essential context for the methodologies presented in Section 3.

Baden-Württemberg (Germany) is a federal state with a budget of approximately €70.2 billion for 2026. It had already implemented a tagging exercise for 2023/24 government expenditure, identifying budget items that contribute positively to climate and environmental goals. Through the project, a complementary methodology was developed and piloted to extend tagging to neutral and harmful spending in the climate domain. The methodology now provides a basis for the state administration to decide its next steps.

Lombardy is Italy’s most populous and economically significant region. It had several building blocks already in place: a Catalogue of Environmentally Harmful Subsidies, a Regional Programme for Sustainable Development, and an Interdepartmental Working Group for the implementation of green budgeting practices. The project drew on these foundations to build a complete tagging methodology from the ground up. Anchored to the EU Taxonomy and covering all classification categories, the methodology provides, for the first time, a shared basis for assessing the environmental relevance of regional spending.

Greece operates a national tagging system covering all six EU Taxonomy objectives across the central government budget, developed progressively since 2020 when it joined the Paris Collaborative for Green Budgeting. The project provided a diagnostic analysis and recommendations for enhancing the existing system. These include options for increasing the granularity of tagging, extending coverage to revenues and tax expenditures, and strengthening the link between tagging results and the budget preparation process.

¹⁴ This section is informed by the diagnostic analyses carried out by project’s technical partners for each Beneficiary Authority at the beginning of the project.

Together, these contexts illustrate three distinct pathways into green budgeting: building a new system from scratch (Lombardy), extending an existing system to cover neutral and negative spending (Baden-Württemberg), and enhancing an established methodology for greater depth and impact (Greece).

Figure 1. Three pathways into green budgeting

ADMINISTRATION	PRE-EXISTING	DEVELOPED/RECOMMENDED UNDER THE PROJECT
BUILD FROM SCRATCH LOMBARDY Region - Italy	<ul style="list-style-type: none"> • Environmentally Harmful Subsidies Catalogue • Sustainable Development Programme • Interdepartmental Working Group 	<ul style="list-style-type: none"> • Complete tagging methodology • EU Taxonomy alignment • 5 classification categories • Decision tree & templates • Pilot across selected missions
EXTEND EXISTING BADEN-WÜRTTEMBERG Federal state - Germany	<ul style="list-style-type: none"> • Positive tagging (since 2023) • Net-zero 2040 target • Green bonds (first issued in 2021) • Sustainable Financial Investment Act (2023) 	<ul style="list-style-type: none"> • Neutral & negative tagging • Adapted brown list • Dual counterfactual method • Pilot across 3 ministries • Finance-environment task force
ENHANCE & DEEPEN GREECE National - Central government	<ul style="list-style-type: none"> • National tagging (since 2020) • 6 EU Taxonomy objectives • Full budget coverage 	<ul style="list-style-type: none"> • Increase tagging granularity • Extend to revenues & tax expenditures • Link to budget preparation • Ex-ante and ex-post impact assessment of policies

3. Methodologies developed and tested

The project supported two Beneficiary Authorities – Lombardy (Italy) and Baden-Württemberg (Germany) – in developing new green budget tagging methodologies. Greece, which already had an operational system covering all six EU Taxonomy objectives, received a diagnostic analysis and recommendations for improving it. This section therefore focuses on the two new methodologies, walking through the five steps of their development. Insights from the Greek experience inform the lessons and recommendations in Section 4.

An illustrative example: Incentives for low-emission biomass heating

To show how green budget tagging works in practice, this section follows one budget item through the methodology: an incentive scheme for replacing highly polluting domestic heating systems with certified low-emission biomass alternatives. At first glance, this looks straightforwardly “green” – it supports the shift away from fossil fuel heating and reduces greenhouse gas emissions. But a closer look raises questions that a tagging methodology must answer:

1 Biomass burning produces particulate matter (PM10, PM2.5) and nitrogen oxides – **does this cause significant harm to the air-quality objective?**

2 **Where does the biomass come from?** Unsustainable forestry or long-distance transport could undermine the climate benefit.

3 **What is the scheme actually replacing?** The climate impact depends on what would happen otherwise – are households switching from oil, gas, or older biomass systems?

This example, drawn from a real item in Lombardy's 2024 budget, shows that “green” is not obvious, even for measures with clear environmental intent. The tagging methodology must provide a structured way to answer these questions consistently.

STEP 1:

Anchoring in policy priorities

What this step is about

Connecting green budget tagging to existing climate and environmental commitments, so that the tagging framework reflects real policy goals rather than abstract categories.

Key questions

What climate and environmental goals has the administration committed to? What laws, strategies, or targets should tagging support? How does GBT fit with existing policy frameworks?

Outputs

A clear link between the tagging framework and the laws, strategies, and targets that define what “green” means in the administration's context.

Green budget tagging does not begin with technical definitions. It begins with fit to policy. Before deciding how to classify expenditures, an administration must identify the commitments and goals that tagging should serve. Anchoring in policy goals ensures that the tagging framework reflects real priorities and supports consistent assessment across budget areas.

Policy anchoring in practice

Both Lombardy and Baden-Württemberg anchored their tagging frameworks in existing laws and strategies.

Lombardy built its tagging approach into the Region's sustainability framework, anchored in the Regional Programme for Sustainable Development (PRSS) – the primary strategic framework for the current legislature, adopted by the newly elected Regional Council in March 2023. The tagging framework draws on multiple sector-specific instruments mapped to each EU Taxonomy objective, including the Regional Climate Law (most recently updated in July 2025) and its key implementation tool for adaptation, the Regional Integrated Strategy for Climate Change Adaptation (SIRACC) adopted in November 2025. Together, these and the other underpinning instruments set binding targets, such as a 43.8% cut in greenhouse gas emissions by 2030, and define the environmental priorities that tagging is designed to track.

Baden-Württemberg designed its tagging framework in line with the objectives of the state's Climate Protection and Climate Change Adaptation Act (KlimaG BW), which requires climate neutrality by 2040 and a 65% emissions cut by 2030 compared to 1990 levels. While broader sustainability frameworks also exist in the state, KlimaG BW is the legally binding instrument with concrete targets, and was therefore the natural reference point. The initial focus on climate mitigation and adaptation also reflected a deliberate effort to keep the additional administrative burden as low as possible while ensuring high-quality outcomes. Reducing bureaucracy is a political priority in Baden-Württemberg. Against this background, it was decided to adopt a phased approach, starting with a limited scope and allowing for future expansion. The other four EU Taxonomy objectives are therefore currently not within scope, but they can be incorporated at a later stage, once Baden-Württemberg has gained further experience and optimised its processes.

The two approaches illustrate a fundamental design choice: comprehensive from the start, or focused with room to expand. Lombardy aligned its framework with multiple sector-specific strategies, each mapped to one of the six EU Taxonomy objectives, building a comprehensive framework from scratch. Baden-Württemberg designed its framework in line with a single climate law, building on an existing system of climate and environmental-positive flags and focusing the new tagging exercise on a well-defined scope. Both choices reflect their institutional starting points.

Applying policy anchoring: the biomass heating example

In Lombardy, the biomass heating scheme connects to PREAC targets on emissions from heating systems and to PRIA goals on air quality - the measure is designed to serve decarbonisation and air-quality objectives together. The DNSH screening checks this alignment, and the scheme's requirements - high-efficiency certified systems with emission-reduction technology - reflect the region's choice to pursue climate action without compromising air quality.

In Baden-Württemberg, a comparable subsidy would be framed through a climate-only lens, asking whether the expenditure is harmful to climate mitigation or adaptation; air pollution, water, and biodiversity impacts could be considered if the methodology were expanded. The same type of spending is assessed differently depending on what the tagging framework has been anchored in.

CHECKLIST: STEP 1 – ANCHORING IN POLICY PRIORITIES

Have you...

- Identified the climate and environmental laws, strategies, and targets that tagging should support?*
- Established a clear link between tagging and these policy frameworks?*
- Decided which environmental objectives to cover at launch and which to defer to later phases?*
- Secured political commitment to GBT – ideally signalled publicly?*
- Communicated the policy rationale to the departments and officials who will be involved?*
- Clarified that tagging assesses environmental impact, not whether spending should continue – results inform decisions, they do not dictate them?*

STEP 2:

Defining the scope

What this step is about

Deciding which parts of the budget will be tagged, against which environmental objectives, and at what level of detail – balancing thoroughness with what is practical to carry out.

Key questions

Which parts of the budget are covered – spending, revenue, or both? Which environmental objectives will be assessed? Which sectors or budget areas are included? How detailed will the tagging be?

Outputs

Clear decisions on what is in scope, what is excluded, and the level at which tagging will happen.

Once the tagging framework is anchored in policy priorities (Step 1), the next decision is scope: which parts of the budget will be tagged, against which environmental goals, and at what level of detail?

Both regions made practical choices to keep the tagging framework workable while still covering the most significant spending. The table summarises the main scope decisions; the two areas that drove the most reflection – pilot design and level of detail – are discussed in the subsections that follow.

Dimension	Lombardy	Baden-Württemberg
Budget coverage	Spending only; revenues excluded	
Environmental objectives	All six EU Taxonomy objectives	Climate mitigation and adaptation only; the other four EU Taxonomy objectives are optional for a later phase
Level of detail	Individual commitment decrees - the administrative acts that create specific spending obligations, and the lowest operational level of the regional budget	Budget items (Titel) - individual expenditure lines within each ministry's budget, the standard level at which funding is authorised and tracked. At earlier stages of the budget cycle, tagging takes place at a less granular level, since fewer details are available at that point
Classification per item	Each item assigned a single tag - no proportional splitting	Items can be split proportionally across multiple tags (e.g., 60% green, 30% grey, 10% brown) - not fixed coefficients like Rio Markers
Exclusions	Staff costs, debt servicing, transfers outside regional control	Staff costs, debt servicing, interest payments, and special financing items centrally tagged as neutral by the task force without case-by-case review - a deliberate choice to reduce the burden on the line ministries

Pilot design

Piloting was carried out in both regions, and the two pilots – documented in dedicated reports – produced practical findings that have already informed the methodologies and shaped several of the lessons in Section 4.

Lombardy designed its pilot around learning rather than coverage. Between October 2025 and January 2026, around 60 commitment decrees were sampled across seven of the region's 18 budget Missions, regardless of spending size, with the aim of testing the methodology across varied policy contexts and engaging departmental managers in its application.

The pilot confirmed that commitment decrees are the right unit of analysis. It also identified the most important operational constraint: decree texts often do not contain enough environmental detail for consistent classification without contextual knowledge held by the responsible manager – an issue returned to in Section 4. An exploratory comparison with AI-assisted tagging on the same 60 decrees was conducted at the end of the pilot to test what additional value AI could bring; results are discussed in Box 2.

Baden-Württemberg ran a coverage-driven pilot from August 2025 to March 2026, tagging approximately 320 items across the 2025/26 biennial budget in three of the state's 17 budget sections – the Ministry of Finance, the Ministry of the Environment, Climate and Energy, and the Ministry of Rural Areas. In the second phase, the pilot also covered around 20 past requests for new funding per ministry (see Box 1). This design produced a complete picture of those sections and tested the methodology at scale, building on the familiarity that line ministries already had with the pre-existing tagging system. The pilot confirmed that the assessment protocols are workable in practice, and surfaced findings that are informing the continued refinement of the methodology. It also confirmed that tagging remains resource-intensive, particularly in the first cycle.

The choice between the two designs depends on how much methodological learning the pilot still needs to absorb. An administration working with a new methodology benefits from a depth-first design; one extending an established approach can afford a broader coverage test. In both cases, the value of piloting lay less in the numbers produced than in the practical insights generated for refining the methodology before full roll-out.

The granularity question

The level of detail at which tagging happens matters for accuracy. International experience shows that tagging at programme level – rather than individual budget items – risks overestimating green spending, because programmes often bundle activities with different environmental impacts. Both regions therefore chose finer-grained approaches, but through different routes: Lombardy through the granularity of the unit tagged (commitment decrees), Baden-Württemberg through proportional allocation within broader budget items (Titel). The blanket central tagging of routine items (personnel, debt servicing, special financing) as neutral complements BW's approach by reducing the number of items requiring individual assessment.

Applying this to the biomass heating incentive:

In Lombardy, the biomass heating incentive falls within Mission 9 (Sustainable development and environmental protection), one of the seven Missions in the pilot. As capital spending with a specific commitment decree, it is in scope and assessed at decree level – meaning the scheme is evaluated as a distinct item with its own classification, against all six EU Taxonomy objectives (climate mitigation as the main objective, with checks against the other five for potential harm). Tagging at this level of detail matters: in a system that classified spending at programme level, the scheme might be bundled with other energy measures, making precise assessment harder.

Under Baden-Württemberg's methodology, a comparable scheme would sit in the budget section of the Ministry of the Environment, Climate and Energy or the Ministry of Rural Areas. It would be assessed against climate mitigation and adaptation only; biodiversity effects of biomass sourcing would not be formally captured at this stage. If the subsidy were part of a broader budget item alongside other measures, the percentage-splitting mechanism would allow the green and non-green components to be distinguished within the same item.

CHECKLIST: STEP 2 - DEFINING THE SCOPE

Have you...

- Decided whether to cover expenditure only, or also revenues and tax expenditures?
- Chosen which environmental objectives to include (all six EU Taxonomy objectives, or a subset)?
- Defined which budget areas or sectors are in scope for the pilot?
- Determined the level of detail at which tagging will occur (programme, budget item, decree)?
- Identified what is explicitly excluded (e.g. personnel costs, debt servicing, transfers)?
- Decided whether items receive a single tag or can be split proportionally across tags?
- Tested the chosen scope against available capacity and data quality?

STEP 3:

Defining what is “green” (and “brown”)

What this step is about

Setting the criteria that determine whether spending counts as environmentally positive, negative, or neutral – the technical core of the methodology.

Key questions

What framework defines environmental contribution (EU Taxonomy, national criteria, or a mix)? What counts as a real contribution to an environmental goal? How is harm to other objectives checked? What counts as environmentally harmful (“brown”)?

Outputs

Criteria for positive contribution to each environmental objective; rules for checking harm to other objectives; a brown list or similar guidance on harmful spending; and the full set of classification categories that result from applying these criteria.

Step 1 identified which policy objectives tagging should serve. This step turns those objectives into operational criteria: what makes a specific expenditure “green,” “brown,” or something in between? Both regions used the EU Taxonomy as a reference, but applied it differently – one comprehensively across all six objectives, the other selectively. The choice of how much of the Taxonomy to operationalise is itself a design decision that needs to be aligned with the policy objectives.

An administration may define all classification criteria at once – as Lombardy did – or build incrementally. Baden-Württemberg took the latter path: having already established criteria for identifying climate and environmental positive spending through a prior tagging exercise, the new methodology focused on defining “brown” and “grey.” What matters is that the resulting framework is complete and internally consistent, regardless of the order in which it was built.

Defining “green” in practice

Lombardy adopted the EU Taxonomy’s two-step logic:

- Substantial contribution: Does the spending make a significant positive contribution to at least one of the six environmental objectives?
- Do No Significant Harm (DNSH): Does it avoid significant negative effects on the other five objectives?

To put this into practice, Lombardy developed detailed criteria for substantial contribution to each objective, along with DNSH screening questions. These draw on EU Taxonomy technical criteria but are simplified for budget tagging. The criteria, the DNSH questions, and a single integrated decision tree were designed together, producing a complete classification framework in one step.

Pilot testing confirmed that the classification categories and substantial contribution criteria were clear and intuitive for assessors. Assessors found the DNSH screening more demanding, particularly in complex or borderline cases, where the questions were found to be overly prescriptive. The methodology is being refined in response: DNSH questions are being simplified, and a shared glossary is being developed to clarify terms such as “impact,” “relevant,” “positive,” and “negative” that different assessors were interpreting differently.

The pilot also surfaced ambiguity between “neutral” (negligible environmental impact) and “untagged” (insufficient information to assess). This distinction proved difficult to apply consistently. The methodology has been adjusted to make information sufficiency an explicit first gate in the decision tree: untagged is now reached only when there is not enough information to assess environmental effects at all, rather than as a residual catch.

Baden-Württemberg already had criteria for identifying climate-positive and environmental-positive spending, established through its tagging system for the 2023/24 budget, and aligned with the logic of the EU Taxonomy. Items assessed as climate-positive under that system automatically carry over as “green” tags – the new methodology does not revisit them. The current exercise therefore concentrates on the remaining items, identifying those that are harmful to the climate and those that are neutral. For spending not already tagged green, the methodology asks two core questions depending on the type of expenditure:

- For funding programmes and grants: Does the spending increase greenhouse gas emissions or worsen climate adaptation compared with the counterfactual – the situation that would prevail if the spending were not made?
- For other state spending: Is the spending compatible with the State’s climate goals? In practice, this is operationalised through a simple test: does the spending lock in the use of fossil fuels?

These questions define what counts as “brown” in Baden-Württemberg’s system. Spending that does not trigger either question – and is not on the brown list – is classified as neutral (“grey”).

Identifying “brown” spending

Both regions developed reference lists to help identify spending that is presumed environmentally harmful.

Baden–Württemberg adapted the European Commission’s list of broadly brown budgetary items to its regional context. This involved narrowing the list to focus on climate mitigation and adaptation, and removing items outside the competence of a subnational government – such as military spending, which falls under federal responsibility. The result is a regional brown list of expenditure types presumed negative for the climate, such as fossil fuel subsidies and gas network expansion. The list continues to evolve through negotiation between the Ministry of Finance and line ministries on borderline cases such as certain agricultural and transport–related spending categories.

Lombardy developed a list of broadly brown expenditures organised by budget Mission and consistent with the EU Taxonomy and the EU GBRF, providing examples of harmful spending for each of the 18 Missions – from fossil fuel incentives in energy policy to subsidies for intensive agriculture. In contrast to Baden–Württemberg, the list does not serve as a fast-track; all items go through the full decision tree.

Classification categories

The criteria defined in this step produce the full set of categories each region uses to classify expenditure. These categories flow directly from the assessment logic: Lombardy’s DNSH step, for example, is what makes it possible to distinguish “green” from “mixed.”

Tagging categories: Lombardy and Baden–Württemberg compared

CATEGORY	LOMBARDY	BADEN-WÜRTTEMBERG
Positive	Green - substantial contribution to at least one objective and passes all DNSH checks	Green - significant positive contribution to climate targets (criteria established in prior flagging exercise)
Positive with risks	Mixed - contributes to one objective but fails one or more DNSH checks	No equivalent - DNSH is not applied, so this distinction does not arise
Negative	Brown - clearly negative impact on one or more environmental objectives	Brown - significantly counteracts climate targets
No significant impact	Neutral - no significant environmental impact relative to the counterfactual	Grey - no discernible significant climate impact
Cannot be assessed	Untagged - insufficient data or no scientific consensus	Not determinable - insufficient information or scientific ambiguity; can only be assigned by the task force

Lombardy’s five categories reflect the breadth of its assessment: six objectives plus DNSH screening produce more possible outcomes. Baden–Württemberg’s four categories reflect both its focused scope and its incremental approach: without DNSH and with only climate objectives in play, the “mixed” distinction does not arise. Both systems include a residual category for items that cannot yet be classified, though Baden–Württemberg restricts its use to the task force – a safeguard against departments parking difficult items to avoid a brown tag.

CHECKLIST: STEP 3 - DEFINING WHAT IS 'GREEN' (AND 'BROWN')

Have you...

- Chosen a reference framework for defining environmental contribution (EU Taxonomy, national criteria, or a hybrid)?
- Developed criteria for what counts as substantial contribution to each environmental objective in scope?
- Established how harm to other objectives will be assessed (e.g. DNSH screening)?
- Developed or adapted a brown list identifying spending types presumed to be harmful?
- Defined the full set of classification categories (e.g., green, brown, mixed, neutral, untagged) and how they relate to the assessment criteria?
- Tested the criteria on sample budget items to confirm they are workable in practice?

STEP 4:

Assessment procedures

What this step is about

Turning the criteria and categories defined in Step 3 into a structured process that assessors can follow consistently – including decision trees, reference scenarios, rules for borderline cases, and documentation requirements.

Key questions

What steps does an assessor follow to classify an item? What reference scenario is used to judge environmental impact? How are borderline or complex cases handled? How are decisions documented?

Outputs

Defined categories with clear meanings; rules for deciding which category applies; guidance on difficult cases.

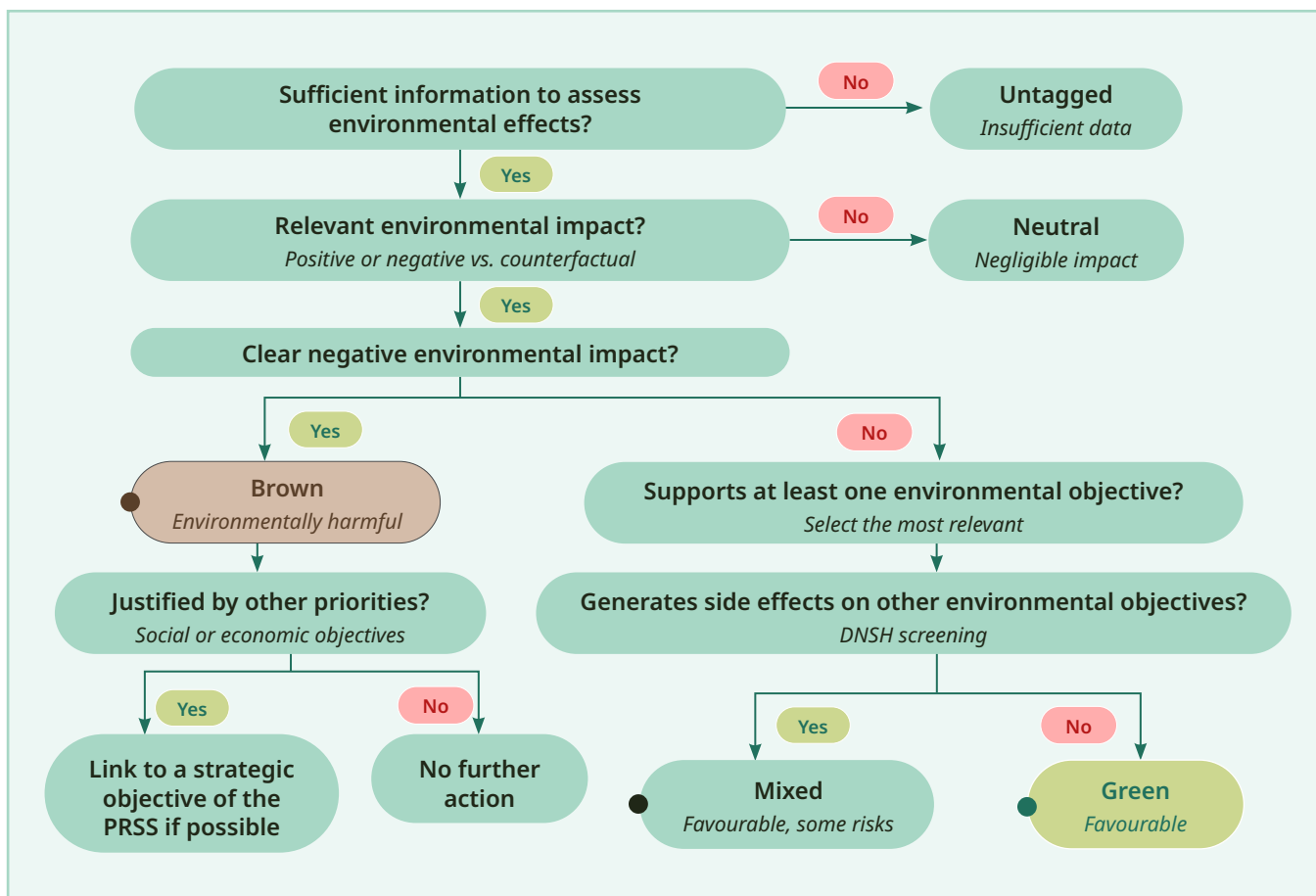
Step 3 defined what counts as green, brown, and neutral. This step sets out how assessors apply those definitions in practice. A well-designed assessment procedure ensures that different people assessing similar items reach consistent results.

Decision structures

The two regions designed different assessment structures, reflecting their different starting points and scope.

Lombardy uses a single integrated decision tree that covers all five classification categories in one process. The assessor works through a sequence of questions – information sufficiency, environmental relevance against a counterfactual, negative impact, positive contribution, and side effects on other objectives – as illustrated below.

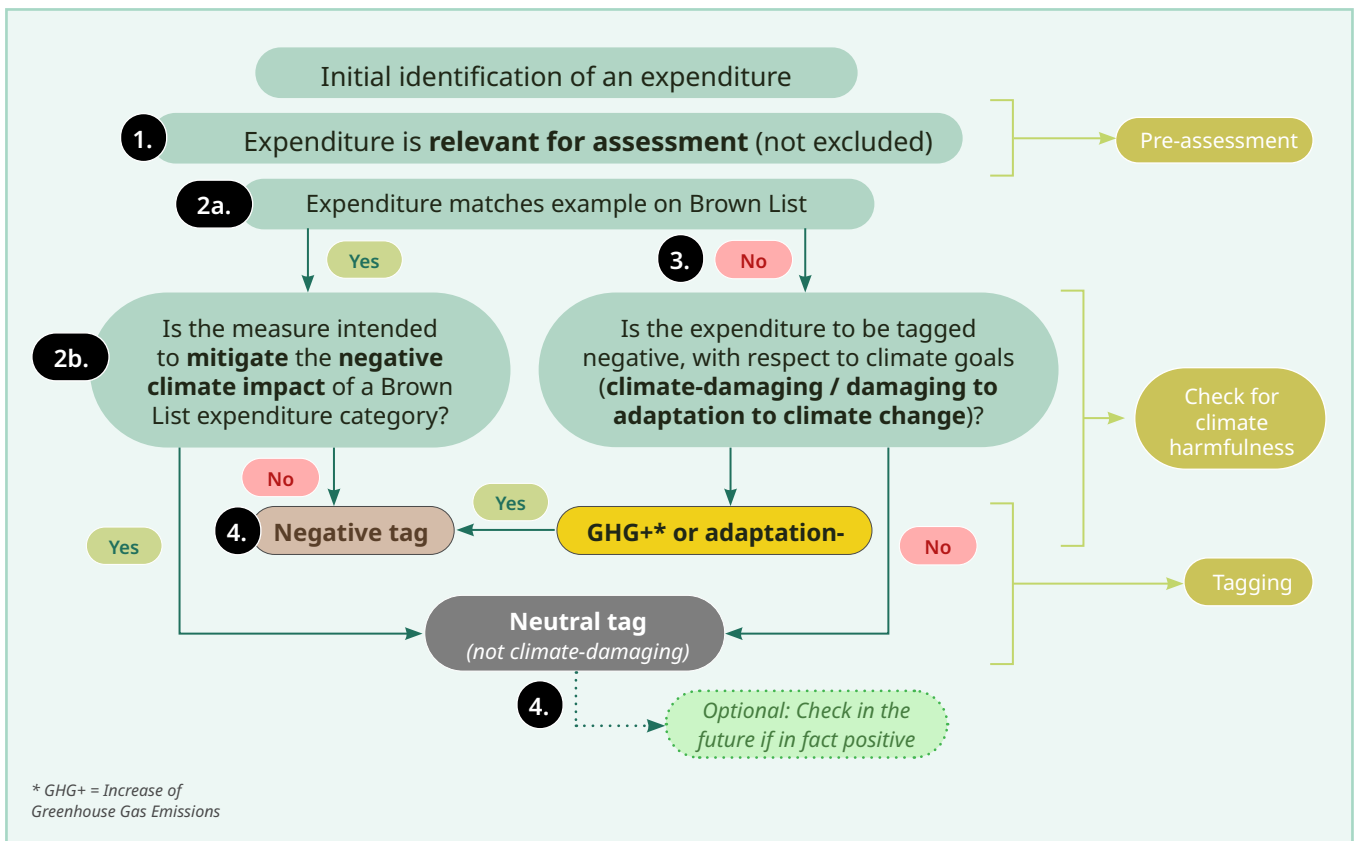
Figure 2: Lombardy tagging decision tree (simplified version)



Pilot testing confirmed that the tree works well in practice. Participants found it easy to apply, and officials were able to use it autonomously after a single training session combining methodological presentation with collaborative classification of real decrees. The sequential logic guided assessors effectively through the classification, even without prior experience of GBT.

Baden-Württemberg uses two separate protocols, both focused on identifying brown and grey items – green tags are carried over from the prior tagging exercise and are not reassessed. Both protocols share the same structure: a check against the Brown List of presumed-harmful expenditure categories, a correction step that asks whether the specific measure is intended to mitigate the climate harm of a brown category, and an individual assessment for items not on the list. The correction step prevents mechanical brown tagging of measures whose explicit purpose is to reduce the climate harm of an otherwise brown sector. The two protocols differ at the individual assessment step: Protocol A applies to funding programmes and grants, asking whether the spending is harmful to climate targets compared to a counterfactual; Protocol B applies to all other state expenditure, asking whether the spending is compatible with the state’s climate goals, operationalised through whether it perpetuates the use of fossil fuels. The pilot confirmed that both protocols are workable in practice, though practical application flagged the need for clearer guidance on some borderline cases.

Figure 3: Baden–Württemberg tagging decision tree (assessment Protocol A, expenditure on funding programmes, and other allocations and grants)



The brown list functions as a fast-track filter at step 2: items that clearly match are tagged brown unless they fall under a correction step that recognises measures explicitly intended to mitigate the climate harm of a brown category. Items tagged grey/neutral may be cross-checked in the future to determine whether they qualify as green.

Reference scenarios

Both regions use counterfactual reasoning to assess environmental impact – comparing the situation with the expenditure against a scenario without it. But they apply this reasoning differently.

Lombardy applies counterfactual reasoning to all expenditure, with the baseline varying by type of spending:

- **Capital expenditure:** the counterfactual is the absence of the intervention (what if the investment were not made?)
- **Current expenditure:** the counterfactual is a reduction in spending (what is the marginal environmental impact of the current level?)

Counterfactual reasoning is treated as a conceptual reference that shapes how assessors answer the decision-tree questions, rather than as an explicit step in the tree itself.

Baden–Württemberg applies counterfactual reasoning only to funding programmes and grants (Protocol

A), where the reference scenario is the status quo without the expenditure. For all other state spending (Protocol B), a counterfactual is not practical – core state functions (such as internal security or policing) often have only one way to be provided, and green budgeting is not intended to question whether those functions should be performed. Protocol B asks whether the spending is compatible with the state’s climate goals, operationalised through a simple test: does the spending perpetuate the use of fossil fuels? This gives assessors a clear, consistent criterion without requiring complex modelling.

Handling partial allocation

When a budget item funds multiple activities with different environmental impacts, the two regions handle the mix differently – and the difference flows directly from the granularity of the tagging unit. Lombardy assigns each item a single tag: tagging at the decree level, where each item typically has a single clear purpose, makes this workable, and items with mixed impacts are captured by the “mixed” category. Baden-Württemberg allows percentage-based splitting – for example, 60% green, 30% grey, 10% brown – because budget items at Titel level often bundle measures with different climate impacts.

BOX 1: Tagging new funding requests in Baden-Württemberg

Baden-Württemberg’s methodology applies tagging in two stages. The first covers the existing state budget; the second covers new funding requests – the list of additional spending proposals that line ministries submit to the Ministry of Finance before each budget cycle, tagged at the level of the individual measure rather than the budget item. The rationale is that tagging results available at the point when new spending is being considered can directly inform negotiations, rather than describing decisions already taken.

The BW pilot tested this approach in early 2026. The three participating ministries applied the tagging methodology to past additional spending / new funding requests, supported by the Task Force.

The experience confirmed the approach is workable. It also surfaced a practical insight: fitting a full tagging assessment into the tight budget planning timeline requires close coordination between the Task Force, the budget department, and the line ministries. Options to explore

include:

- Lighter-touch assessment formats for the submission stage;
- Early engagement with line ministries so the climate dimension is considered during drafting and the subsequent submission of new funding requests;
- Phased integration starting with a subset of new funding requests.

Lombardy shares this ambition: having begun with ex-post tagging to build internal acceptance, the region intends to explore extending tagging upstream into the budget preparation process in a next phase.

Ex-ante tagging is where tagging has the greatest potential to influence budget decisions – and where the practical challenges are most concentrated. Baden-Württemberg’s experience offers one of the first tested models for making it work.

Applying this to the biomass heating incentive:

Under Lombardy's decision tree, the biomass heating incentive is assessed in a single pass through the full tree:

ASSESSMENT STEP	RESULT	REASONING
Sufficient information to assess environmental effects?	Yes	The decree was detailed enough for the assessment.
Relevant environmental impact?	Yes	The scheme affects greenhouse gas emissions from residential heating
Clearly negative environmental impact?	No	It reduces emissions by replacing older, higher-emission heating systems
Supports at least one environmental objective?	CCM	Climate change mitigation
Generates side effects on other environmental objectives? (DNSH screening)		
Climate adaptation	No	No climate risk from the intervention
Water	No	No water use or discharge involved
Circular economy	No	Decree requires controlled disposal of old systems
Pollution	No	Systems must have top environmental certification and emission filters (PM10, PM2.5, NOx)
Biodiversity	No	Biomass must be certified and locally sourced from sustainable supply chains
Classification: Green – Substantial contribution to climate mitigation with full DNSH compliance		

Under Baden-Württemberg's new tagging methodology, the same type of spending would be assessed using the criteria for funding programmes. The methodology asks whether the spending increases greenhouse gas emissions or worsens climate adaptation compared to the status quo. Since the scheme reduces emissions rather than increasing them, it would not be tagged brown. In practice, such a

scheme would most likely already carry a green tag from Baden-Württemberg's prior tagging exercise and would be carried forward unchanged. For a similar scheme not yet tagged green, the new methodology would classify it grey (neutral) at this stage, with a note to cross-check in a later round whether it qualifies as green.

ASSESSMENT STEP	RESULT	REASONING
Relevant for tagging?	Yes	Not excluded or covered by blanket tagging
Matches the Brown List?	No	A heating replacement scheme shifting households to lower-emission systems does not match any Brown List category
INDIVIDUAL ASSESSMENT (PROTOCOL A)	RESULT	REASONING
Does it increase GHG emissions or worsen climate adaptation vs. the status quo?	No	It replaces older, higher-emission heating systems - emissions decrease rather than increase
■ Classification: Grey - Not harmful to the climate (neutral tag)		

The comparison shows how scope choices translate into different depths of assessment for the same expenditure. Lombardy's DNSH screening formally weighs pollution and biodiversity effects; Baden-Württemberg's current

methodology does not. The climate impact is captured either way, but a scheme receiving a green tag in Baden-Württemberg today may receive a different tag when the methodology extends to further Taxonomy objectives.

Information requirements

Both methodologies require assessors to record a written justification for each classification decision. This serves several purposes: it supports consistency, enables quality checks and audits, creates a training resource for future rounds, and provides a foundation for digital support tools. Consistent documentation also depends on assessors having access to the right information in the first place. For the biomass heating incentive, for example, the assessment drew on:

Information needed	Where it comes from	What it helps determine
Scheme goals and eligibility rules	Decree text	Check environmental intent and safeguards
Technical requirements for funded systems	Decree text / programme guidelines	Check emission standards, certification requirements
Biomass sourcing requirements	Decree text / programme guidelines	Check sustainability of fuel supply
What systems are being replaced	Programme rationale / sector context	Assess emissions impact relative to the counterfactual
Disposal requirements for old systems	Decree text	Check circular economy compliance

In Lombardy, the decree text provided enough detail to complete the assessment. Where decrees are shorter or refer to inaccessible annexes, the assessor draws on the upstream policy documents the decree implements and on the manager's contextual knowledge of the spending. The residual "untagged" category is reserved for cases where these sources do not together provide enough information to assess environmental effects.

Baden-Württemberg grounds this in three principles: objectivity, a holistic approach, and verifiability – documentation must be based on expert information, reflect full lifecycle effects, and be checkable by third parties.

BOX 2: Can AI support green budget tagging?

Lombardy's pilot included an exploratory comparison between manual tagging by 80 regional officials and AI-assisted classification using the same decision tree logic. The AI (ChatGPT Edu) received only the decree text and committed amount – no contextual knowledge, no upstream policy documents, no environmental assessments. The results showed 57% full convergence, confirming that the methodology produces consistent outcomes when sufficient information is available in the administrative act. Most of the remaining 37% were caused by information gaps: officials could draw on institutional knowledge not documented in the decree, while the AI classified these items as untaggable.

In a small number of borderline cases, the AI applied the methodology more cautiously than human assessors – flagging potential environmental risks where officials had given the benefit of the doubt. These divergences can be read two ways. They may signal a limited risk of optimism bias in manual classifications, where officials tend to favour green outcomes when explicit safeguards are absent. But they may also reflect that only officials responsible for the decree have access to the upstream policy intent and the manager's contextual knowledge of the activity, which the AI, working only from the decree text, cannot draw on.

The comparison reinforces two practical lessons. First, administrative documentation needs to improve if tagging is to be scalable and consistent. Second, AI-assisted tools show genuine promise as a complement to – not a replacement for – expert judgment. The 2025 European Commission Survey found that no Member State respondent was using AI tools for tagging, making the Lombardy's comparison a notable early example in the field. Interest in AI support is increasingly shared across administrations, however: the effort of establishing and running a tagging system is commonly underestimated, and the manual burden of each cycle remains substantial even where methodologies are mature.

CHECKLIST: STEP 4 – ASSESSMENT PROCEDURES

Have you...

- Developed a step-by-step assessment procedure (decision tree, protocol, or similar) that assessors can follow?*
- Defined the reference scenario used to judge environmental impact (e.g., absence of expenditure, reduction in spending, compatibility with policy goals)?*
- Established whether the same procedure applies to all expenditure types, or whether different types require different assessment approaches?*
- Decided whether items can be split proportionally across categories, or whether each item receives a single tag?*
- Provided guidance on how borderline or difficult cases should be handled?*
- Established documentation requirements — what justification must be recorded for each classification decision, and how records will be reviewed and verified?*
- Tested the procedure on sample items to check it produces consistent results across different assessors?*
- Identified what information assessors need to classify each item (e.g. programme objectives, eligibility criteria, technical specifications)?*
- Created reporting templates that capture the classification, reasoning, and supporting information?*

STEP 5:

Establishing governance arrangements

What this step is about

Assigning roles and responsibilities so that tagging is done consistently, reviewed properly, and trusted by stakeholders.

Key questions

Who does the tagging – a central team or individual ministries? Who checks and approves the results? Who decides disputed or borderline cases? Who controls the “untagged” category? How is quality maintained over time?

Outputs

Clear roles for tagging, review, and oversight; ways of coordinating across departments; procedures for quality assurance.

Governance is central to both tagging frameworks. Green budget tagging involves expert judgment, and different assessors can reach different conclusions on the same item. Beyond the templates, guidance and training that support individual assessments, a clear governance framework keeps classification consistent across departments, protects the integrity of results, and manages the risk of managers classifying their own spending too favourably (“greenwashing”).”

Governance models compared

Dimension	Lombardy	Baden-Württemberg
Who tags	Departmental managers (dirigenti) responsible for each spending item	Ministry officials responsible for their budget section
Coordination body	Interdepartmental Working Group	Interministerial Task Force
External support	EU-funded current project, with technical support from SDA Bocconi (methodology development, piloting, training)	EU-funded current project, with technical support from FiFo Cologne (methodology development, piloting, training)
Quality assurance	Working Group ensures consistency across directorates; sample checks and cross-departmental workshops planned	Task Force reviews results
Control of "untagged"	Working Group oversight; "untagged" used only where data genuinely missing	Task Force only – ministries cannot assign "not determinable" themselves
Dispute resolution	Working Group facilitates dialogue; external technical support available	Task Force seeks to reach a decision in dialogue with line ministries

Both regions adopted a decentralised tagging model – officials closest to the spending do the initial classification, with central coordination to ensure consistency. This reflects the informational reality: sectoral entities hold the contextual knowledge about their own spending that a central body cannot easily replicate. Central oversight manages the corresponding risk of inconsistent or self-serving classification.

Control of the “untagged” or “not determinable” category is especially important. Without central control, the residual category risks becoming a means of avoiding difficult classifications. Both regions addressed this by giving the central coordination body control over this category.

Applying this to the biomass heating incentive:

In Lombardy, the biomass heating incentive falls under the Directorate-General for Environment and Climate; the responsible departmental manager does the tagging, and the Interdepartmental Working Group reviews for consistency. In Baden-Württemberg, the equivalent scheme would be tagged by officials in the Ministry of the Environment, Climate and Energy or the Ministry of Rural Areas, with Task Force review. Governance matters most in borderline

cases. In Lombardy, a contested DNSH assessment on air quality would ultimately be decided by the Working Group; in Baden-Württemberg, a disagreement on whether a heating subsidy belongs on the brown list would be resolved by the Task Force. In both systems, resolution of such cases must be documented so that decisions can guide similar cases later.

CHECKLIST: STEP 5 – ESTABLISHING GOVERNANCE ARRANGEMENTS

Have you...

- Assigned responsibility for tagging to officials with knowledge of the spending?*
- Established a central coordination body with authority to review, query, and if necessary, override classifications?*
- Defined who controls the 'untagged' or 'not determinable' category?*
- Set up a process for resolving disputes or borderline cases?*
- Arranged for training and ongoing support for officials doing the tagging?*
- Planned how tagging results will feed into budget decisions – not just reporting?*
- Considered how quality will be assured over time (sample checks, peer review, external assessment)?*

Governance does not end with the initial design. Training for new staff, periodic refresher sessions, and updating of reference lists (in particular the brown list) will need to be embedded into the routine work of the coordination body. Without this, the methodology risks drifting as personnel turn over and new types of spending emerge over time.

The next section draws on the project's experience and peer exchanges with other administrations to identify what worked, what proved difficult, and how to avoid the most common pitfalls.

4. Lessons and recommendations

This section brings together what the project learned and the practical guidance that flows from it. The lessons draw on the experience of three administrations and should be read as practical insights rather than universal rules. Each lesson – drawn from the methodology development, the Lombardy and Baden–Württemberg pilots, the Greek diagnostic, and peer exchanges – is paired with an action-oriented recommendation for administrations considering or refining their own approach. The eight lessons are grouped in three phases: getting started, designing the methodology, and governance and implementation.

GETTING STARTED

1. Anchor green budgeting politically or legally

Green budget tagging needs to be linked to recognised political, legal or strategic commitments to endure. Baden–Württemberg designed its tagging framework in line with the state’s climate objectives, drawing on the framework set out in the Climate Protection and Climate Change Adaptation Act; Lombardy linked it to the region’s broader sustainability commitments; Greece embedded it in ongoing national fiscal reform. Anchoring gives the exercise legitimacy and helps coordination bodies engage sectoral entities.

Anchoring should define the mandate, not predetermine the method. Decisions on scope, ex-ante versus ex-post, or single versus proportional tagging depend on capacity, budget process and intended use of results – not on the legal mandate itself. The Paris and Strasbourg exchanges reinforced both points: legal obligation can create a strong central mandate (France) and national legislation can give impetus to subnational implementation (Grand Est), but in more decentralised systems such as Germany, political anchoring still needs to be complemented by early sectoral entities’ involvement and consensus-building around the tagging framework.

In practice

- Link tagging explicitly to climate laws, fiscal reform commitments or environmental strategies; embed the requirement in legislation or medium-term fiscal frameworks where possible.
- Keep these decisions – scope, ex-ante versus ex-post, single versus proportional tagging – as explicit design choices, not automatic consequences of the mandate.
- In decentralised systems, pair political or legal anchoring with early line-ministry engagement and consensus-building on the methodology.

2. Calibrate scope to context – and be clear about the trade-offs

There is no single right answer to how broad a tagging methodology should be at launch, and both regions made legitimate but different choices. Baden–Württemberg focused first on climate mitigation and adaptation, deliberately limiting scope to manage capacity, build on its existing tagging system, and keep the additional administrative burden as low as possible.

This kept the initial workload manageable but reduced the ability to show mixed effects: when only climate mitigation and adaptation is assessed, a negative tag stays negative, whereas broader coverage may reveal positive or neutral effects against other EU Taxonomy objectives such as biodiversity or water.

Lombardy chose to cover all six EU Taxonomy objectives from the outset. Its technical experts argued that full scope required only marginally more effort than a partial approach and avoided substantial rework later. The caution here is important: expanding scope is not always a simple add-on, since new objectives may require new criteria, guidance, training and ways of handling mixed effects.

In practice

- Decide scope explicitly on three dimensions: which environmental objectives to cover, which parts of the budget to include, and at what level of detail.
- Anchor to the EU Taxonomy where relevant, using it as a reference framework.
- Where only some objectives are covered initially, recognise that adding further objectives later will require additional methodological work - new criteria, guidance, training, and ways of handling mixed effects - not simply an extension of existing rules.

3. Treat early implementation as a learning phase

Across the project, the most valuable early output was not only the tagging results themselves, but what the work revealed about methodologies, constraints and next steps. Lombardy's pilot surfaced ambiguity in the neutral/untagged distinction and overly prescriptive DNSH questions; Baden-Württemberg's pilot showed that budget planning timelines require lighter-touch assessment formats; Greece's diagnostic identified increasing coverage of tagging to include revenue and tax expenditures as priority areas for further work.

Framing early cycles as learning rather than performance matters politically as much as technically. Sectoral entities are more willing to engage with a new methodology – and to accept brown or mixed tags – when they understand that the exercise is being refined in good faith rather than used to judge them.

In practice

- Communicate clearly that early tagging results are intended to improve the methodology and inform future decisions – not to rank departments or judge the quality of their budgets.
- Build feedback loops into piloting, training and review: collect structured feedback from assessors and revise tools before scaling up.
- Plan buffer time – real pilots often take longer than scheduled, and finding problems late is more costly than finding them early.

DESIGNING THE METHODOLOGY

4. Invest in tools, piloting and training together

A methodology that looks complete on paper does not automatically translate into consistent classification in practice. Three elements together determine whether non-specialist assessors can apply it reliably: practical tools that translate definitions into operational steps, piloting that unearths problems before wider roll-out, and training that equips officials to apply the methodology with confidence.

Both regions invested substantially in tools – decision trees, brown lists, DNSH screening questions, reporting templates, worked examples. Documentation templates in particular required significant methodology work: well-designed templates keep assessors aligned on what to record and make the exercise easier to apply consistently. Lombardy's pilot showed that the way questions are framed matters: negatively framed questions such as “Does this measure increase CO₂ emissions?” proved easier for assessors than questions requiring proactive proof of non-harm. A shared glossary is also being developed to stabilise interpretation of key terms. Training mattered equally: a single session combining methodological presentation with collaborative classification of real decrees was enough for Lombardy officials to apply the decision tree autonomously. Two cautions follow – tools should be designed for the local context rather than copied, and the initial development effort should not be underestimated.

In practice

- Develop decision trees, reference lists, templates, worked examples and a shared glossary together – not in isolation.
- Frame screening questions concretely; negatively framed questions (“Does this increase emissions?”) proved easier to apply than requiring proactive proof of non-harm.
- Test tools on a diverse sample and revise them in response to pilot findings before wider roll-out.
- Combine methodological instruction with hands-on classification of real budget items, and involve future users in developing the tools rather than presenting them with a finished product.
- Recognise that first-time methodology development requires substantial time, coordination and technical work.

5. Choose a granular unit – and involve the people closest to the spending

Tagging at a sufficiently granular level is essential for credible results. Programme-level tagging can overstate environmental relevance when only some activities within a programme are genuinely green. Lombardy addressed this through the granularity of the unit tagged (commitment decrees); Baden-Württemberg through proportional allocation within broader budget items. Greece's diagnostic reinforced the same point: increasing granularity was identified as a key consideration for further methodology development.

Greater granularity increases the information required for assessment, and decree or item texts do not always contain all the data needed. Lombardy's AI-assisted comparison illustrated this directly: when the same decrees were classified using only the decree text, AI classifications fully converged with manual

results in 57% of cases, with most divergences reflecting contextual knowledge held by officials but not documented in the decree. The practical implication is that granular tagging works when the officials closest to the spending – who hold that contextual knowledge – are involved in classification, supported by central quality assurance. Where documentation gaps remain, administrative acts should gradually be strengthened to include minimum information on environmental objectives, expected effects and safeguards.

In practice

- Tag at a granular level to avoid overestimations - such as commitment decrees or individual budget items, not programmes.
- Assign first-level classification to line managers or their staff, supported by central quality assurance.
- Treat “untagged” or “not determinable” as a legitimate outcome where information is genuinely insufficient
- Over time, strengthen administrative documentation so that acts include minimum information on environmental objectives, expected effects and safeguards - a prerequisite for future AI-assisted support.
- Where available, draw on existing environmental assessments, such as environmental impact assessment or ERDF criteria, as supporting evidence.

6. Build the methodology to handle negative and mixed effects credibly

Classifying spending as environmentally harmful was the most politically sensitive element of the methodology in both subnational administrations. The challenge has both a methodological dimension – reliable criteria for identifying harm – and a political dimension that may be more important than initially anticipated. Both regions developed tools to reduce reliance on individual judgement: Baden-Württemberg adapted the European Commission’s brown list as a fast-track filter, and Lombardy developed a reference list of broadly brown expenditures organised by budget Mission. In both cases, the aim was to make brown classifications follow from transparent rules rather than contested case-by-case decisions.

Tools alone do not resolve the political challenge. Baden-Württemberg’s pilot showed reluctance in the assignment of negative tags across participating ministries; Lombardy’s experience showed similar sensitivities, with managers and political leaders concerned about how brown percentages would be perceived externally. Several approaches help manage this. Terminology matters – some administrations may prefer softer language than “negative” or “brown”, such as “not aligned”.¹⁵

Broader environmental coverage – including a ‘mixed’ tag, as in the case of Greece – can help address this challenge, since an item negative for one objective may be neutral or positive for another. Framing brown spending in wider policy context rather than as an accusation can also be helpful: Lombardy linked tagging to its regional sustainable development programme, allowing managers to explain environmentally brown expenditure through wider economic or social objectives. The Strasbourg exchange with Grand Est reinforced this point: negative tagging is easier to accept when paired with improvement narratives and transition pathways.

¹⁵ This term is used by New York City in its climate budgeting exercise, see <https://www.nyc.gov/assets/omb/downloads/pdf/exec24-nycbta.pdf>.

In practice

- Develop structured ways to identify harmful spending - adapted brown lists, DNSH-based screening or equivalent methods.
- Where multiple environmental objectives are covered, consider a “mixed” category for spending that is positive on one objective but harmful on another; this can be more politically acceptable than a straight brown tag.
- Consider whether terminology matters in the administration’s context - some may prefer softer alternatives to “negative” or “brown”.
- Frame brown spending in wider context rather than as an accusation, linking it to economic or social sustainability considerations where relevant and pairing it with improvement narratives or transition programmes.
- For new funding requests, consider using DNSH rather than brown lists to bypass some of the political difficulty of classifying existing spending.

GOVERNANCE AND IMPLEMENTATION

7. Combine decentralised tagging with strong central quality assurance

All three frameworks adopted the same basic model: officials familiar with the activities assign the tag, while a central body coordinates, reviews and, where necessary, adjusts results. In Baden–Württemberg, the Task Force established for the pilot brought together the Ministry of Finance and the Ministry for the Environment; in Lombardy, the Interdepartmental Working Group is led by the Directorate–General for Environment and Climate and brings together the Presidency Directorate, the Agriculture Directorate, and the Territory and Biodiversity Directorate; in Greece, line ministries submit their tags to an Interministerial Committee (IMC) for review and quality assurance. Decentralised execution is necessary because sectoral entities hold contextual information that a central body cannot easily replicate; central oversight is essential to ensure consistency and prevent difficult cases from being avoided.

The same logic shapes how the systems handle cases of doubt. When a sectoral entity does not have enough information to assign a tag – or when there is genuine disagreement on how to classify an item – the case is referred to the central body, which decides whether to assign a tag, request additional information, or apply the residual category (“not determinable” in Baden–Württemberg, “untagged” in Lombardy). Restricting these categories to the central body prevents them from becoming a means of avoiding difficult classifications.

Governance design should fit the institutional context. Lombardy’s broader working group helped spread ownership across multiple departments. Baden–Württemberg’s leaner Task Force offered agility, supported by “mirror units” (Spiegelreferate) in the Ministry of Finance for plausibility checks. In Greece, the IMC draws together expertise from the Ministry of Economy and Finance and the Ministry of Environment and Energy, with constant cooperation between the line ministries and the IMC to provide robust evidence for tags assigned. In general, what matters is not whether the group is broad or lean, but whether it has the mandate, capacity and authority to ensure credible results.

In practice

- Establish a central task force or working group with a clear mandate to issue guidance, query classifications and review results.
- Choose a structure that works in context - there is no single right size; what matters is mandate and capacity, not headcount.
- Use existing institutional structures, such as Baden-Württemberg's "mirror units" (Spiegelreferate), for informed plausibility checks.
- Restrict residual categories ("untagged", "not determinable") to the central body, and establish a clear referral process so that sectoral entities know when and how to escalate cases of doubt.
- Require written justifications for every classification decision, creating an auditable record that supports consistency, future scrutiny, training and digital support tools.
- As the system matures, consider independent expert review of the methodology and results to strengthen credibility beyond the executive branch.

8. Connect tagging to budget decisions – not only reporting

Tagging results do not influence decisions automatically. The project's experience shows that administrations need to decide from the outset how tagging information will enter the budget process: through budget instructions, new funding requests, internal dashboards, parliamentary annexes, management discussions or public reporting. Across the project, this connection remains the main area for further development. Greece already includes tagging results in the draft budget submitted to parliament, but its diagnostic identified the need to strengthen the link with budget preparation and proposed that harmful expenditures to be prioritised for reform are highlighted in the budget instructions. Baden-Württemberg designed its methodology with an ex-ante ambition, testing tagging for new funding requests where results could influence decisions before allocations are finalised – though the pilot also showed that tight budget deadlines may require lighter assessment formats. Lombardy started with ex-post tagging to build acceptance and aims to use results increasingly as a management tool for future spending decisions.

Peer exchanges showed several possible routes. France presents its green budget as an annex to the annual finance bill, supporting parliamentary debate, and its inter-ministerial task force meets annually with civil society organisations to review feedback. Grand Est has used public green budgeting documents to build understanding among elected officials over several years. The Basque Country uses tagging to monitor progress against a legally mandated climate-spending target. There is no single correct model; what matters is that results are channelled to the people and processes that shape budget choices. Expectations should also be realistic: green budget tagging tracks the environmental alignment of spending but does not by itself measure actual environmental impact or replace other policy tools. Its value is greatest when it becomes part of a wider cycle of planning, budgeting, monitoring and programme redesign.

In practice

- Make tagging information available at the points where it can influence decisions - in budget instructions, during the draft budget stage, in new funding requests and in management review discussions.
- Choose appropriate channels for different audiences: parliamentary annexes and committee meetings for political engagement; internal dashboards and management dialogues for operational use; public reporting as the system matures.
- When an activity is at risk of being assigned a brown or mixed tag, encourage departments to consider whether design changes - eligibility criteria, technical requirements or environmental safeguards - could improve its classification.
- Where possible, link ex-ante and ex-post tagging to create a continuous planning-monitoring cycle.
- As the system matures, engage external stakeholders or independent reviewers to strengthen credibility.

Annexes

Annex A. About the project

The project “Developing Green Budgeting Practices in Baden–Württemberg, Lombardy, and Greece” supported three Beneficiary Authorities in designing, developing, and implementing green reforms that integrate climate objectives into the budget cycle and redirect public spending toward green priorities. The project ran from September 2024 to June 2026. It is funded by the European Union via the Technical Support Instrument (TSI), implemented by Expertise France, and coordinated by SG REFORM of the European Commission.

Beneficiary Authorities

GREECE	LOMBARDY, ITALY	BADEN–WÜRTTEMBERG, GERMANY
Ministry of Economy and Finance	Ministry of Environment and Climate	Ministry of Finance

Context

The project built on the European Commission’s efforts to help Member States align public financial management with climate objectives. It follows the EU Green Budgeting Training Project¹⁶, which supported 23 Member States in implementing green budgeting practices through training and peer exchanges. This new phase extended support to the subnational level, providing tailored technical assistance for developing green budget tagging methodologies in Baden–Württemberg and Lombardy, while also supporting Greece in refining its existing national approach. The project continued to foster training and peer learning across all three beneficiaries.

The project aligned with the European Green Deal’s goal of making Europe the first climate–neutral continent by 2050. Green budgeting tools are central to this transition, helping to evaluate the environmental impact of budgetary policies, redirect resources toward sustainable priorities, and phase out harmful subsidies.

At the core of the project was green budget tagging – the practice of assessing budgetary items (expenditure, revenue, tax expenditure) and systematically identifying whether they contribute to or impede the achievement of climate and environmental objectives.

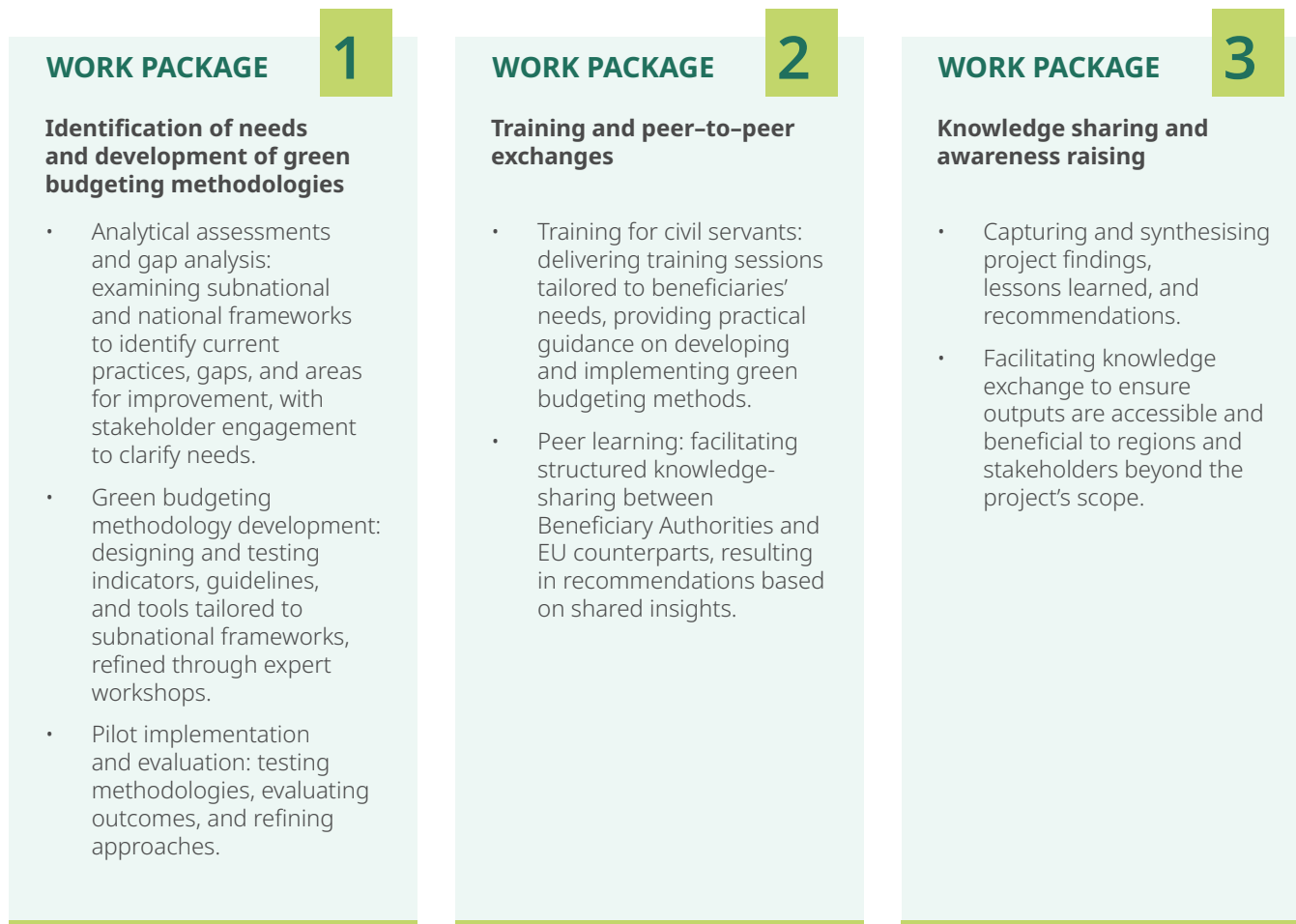
Expected outcomes

01	Improved methodological tools Develop methodological tools to improve accountability and transparency in assessing the coherence of subnational and national climate commitments and strategies.
02	Strengthened capacity Strengthen knowledge and skills among civil servants who operationalise tagging and other green budgeting practices, addressing gaps in skills and capacity.
03	Increased awareness Increase awareness of green budgeting practices among institutional and political actors in Beneficiary Authorities and EU subnational and central administrations beyond the project.

¹⁶ https://reform-support.ec.europa.eu/publications-0/key-achievements-and-success-stories-tsi-green-budgeting_en

Project activities

Technical assistance activities were organised in three interconnected work packages:



Annex B. Further reading

- Botta, F., Bender, C., Lenk, T. and Hesse, M. (2024). Budgetary Framework Reform in Green Transition in the EU and Germany. *Intereconomics*, 59(1), 48–54. <https://doi.org/10.2478/ie-2024-0010>
- Boutron, C. (2023). *Greener, Better, Stronger: Factors for the Successful Implementation of Green Budgeting in EU Member States*. I4CE, Paris.
- Bova, E. (2021). *Green Budgeting Practices in the EU: A First Review*. European Economy Discussion Paper 140. European Commission, Brussels.
- Bagnoli, V. and Cook, J. (2025). *Transforming State and Regional Budgets for Net Zero*. Next Generation Budgets Policy Report, Under2 Coalition.
- Baur, M., Bruchez, P.-A. and Nicol, S. (2021). *Climate Change and Long-term Fiscal Sustainability: Scoping Paper*. GOV/PGC/SBO(2021)4. OECD Working Party of Senior Budget Officials, Paris.
- Department of Finance, Financial Services Division (2025). *Green Budgeting in Ireland from a Tax Perspective – Update to the Methodology*. Dublin: finance.gov.ie.
- European Commission (2021). *Green Budgeting: Towards Common Principles*. doi:10.2765/51675. Brussels.
- European Commission (2023). *Green Budgeting in the EU: Key Insights from the 2023 European Commission Survey of Green Budgeting Practices*. Directorate-General for Economic and Financial Affairs, Brussels, 17 May 2023.
- European Commission (2025). *Green Budgeting in the EU: Key Insights from the 2025 European Commission Survey on Green Budgeting*. Directorate-General for Economic and Financial Affairs, Brussels, 18 September 2025.
- Darvas, Z., Sekut, K. (2025). *Greening the EU Budget: Why Climate Mainstreaming Needs Reform*. Bruegel Analysis, 26 February 2025. <https://www.bruegel.org/analysis/greening-eu-budget-why-climate-mainstreaming-needs-reform>
- Marinheiro, C., Sousa, A. and Pinheiro, A. (2022). *The Climate Dimension of Fiscal Policy Sustainability: Best Practices in Green Budgeting and Lessons for Portugal*. Working Paper No. 01/2022. Portuguese Public Finance Council (CFP), Lisbon.
- OECD (2021). *Green Budget Tagging: Introductory Guidance & Principles*. OECD Publishing, Paris. <https://doi.org/10.1787/fe7bfcc4-en>
- OECD (2022). *Aligning Regional and Local Budgets with Green Objectives: Subnational Green Budgeting Practices and Guidelines*. OECD Multi-level Governance Studies. OECD Publishing, Paris. <https://doi.org/10.1787/93b4036f-en>
- OECD (2022). *OECD Journal on Budgeting, Volume 2022, Issue 2*. OECD Publishing, Paris. <https://doi.org/10.1787/e54588a3-en>
- OECD (2024). *Beyond Green Tagging: Linking Budgets Better with Climate Goals*. GOV/SBO(2024)15. Paper presented at the 46th Annual Meeting of Senior Budget Officials, Paris, 3–4 June 2024.

- OECD (2024). Green Budgeting in OECD Countries 2024. OECD Publishing, Paris. <https://doi.org/10.1787/9aea61f0-en>
- Pojar, S. (2022). Environmental Assessments within Green Budgeting. European Economy Discussion Paper 175. European Commission, Brussels.
- Pojar, S. (2023). How Green Budgeting is Embedded in National Budget Processes. European Economy Discussion Paper 196. European Commission, Brussels.
- Pojar, S. (2024). Green Budgeting in the EU – Using Green Budgeting to Strengthen the Alignment of Public Financial Flows with Climate and Environmental Objectives. Presentation at the HP4 Workshop, 3 June 2024. European Commission, DG ECFIN.
- Pojar, S. and Bärnreuther, J. (2025). Green Budgeting in the EU: How to Align the Various Green Workstreams for a Coherent Approach and Improved Policy Making? European Economy Discussion Paper 230. European Commission, Brussels.
- Salmon-Genel, M. (2025). Assessing the Macro-Fiscal Risks from Climate Change: Concepts, Methodological Approaches, and Insights from Country Practices. European Economy Discussion Paper 224. European Commission, Brussels.